Modelling and comparing maps with graphs

Florence Le Ber, Jean-Luc Metzger, Amedeo Napoli

Modelling and comparing maps with graphs

Florence Le Ber, Jean-Luc Metzger, and Amedeo Napoli
LORIA UMR 7503, BP 239, 54506 Vandœuvre-lès-Nancy, France
Phone: 33 3 83 59 20 00, Fax: 33 3 83 41 30 79
E-mail: {leber,metzger,napoli}@loria.fr

Abstract: This paper is concerned with the modelling of maps with conceptual graphs for the design of a knowledge-based system. We first describe the maps, which are synthetic descriptions of farm territories. We then explain the modelling principles we have used: the spatial objects of the map and the relations are represented into concepts linked with arcs. The reasoning principle of the system are briefly described. We finally focus on an example of graph comparison.

Keywords: graphs, similarity measure, hierarchies of spatial relations, spatial structures, agronomy.

Introduction

In a previous work, we have used topological relations to recognize landscape patterns on satellite images [8, 7]. We are now interested in the qualitative description of farm spatial organizations. We presently work with agronomists to develop a knowledge-based system that could help them analyze the relationship between farm spatial organization and functioning. Understanding this relationship is important for the good management of several environmental problems like water pollution, soil erosion or landscape changes. Thus, the agronomists use special maps of farm territories, called *choremes*[1], that express both the spatial organization and the functioning of the farms considered [6]. We have decided to use graphs for modelling and comparing these maps. Our aim is to build a case-based explanation system to help analyze a farm spatial organization and functioning from previously analyzed cases.

This paper is organized as follows. The first part describes the maps and the way they are modelled with graphs. The second part gives the reasoning principles used in our system. The third part focuses on the computation of similarity between graphs.

1 Modelling maps with graphs

An example of a choreme which represents the spatial organization and functioning of a farm in Lorraine is shown in Figure 1. Several objects are represented: a town, a village, a river, a wood, two roads, pastures and crop fields. Icons are used to point out particular situations or roles of the objects. Agronomists interpret the spatial organization of these objects with respect to the farm functioning. For example, they say that the town is an obstacle for the farmer (since it separates the farm territory in two parts) whereas the village is not, or that the farmer uses meadow to isolate crop fields from river and wood for protecting crops from dampness and shade.

We have chosen graphs for modelling the spatial organization of a farm. Graphs are well-suited for representing complex real-world objects and for communicating with domain experts. The graphs we use are inspired from conceptual graphs [3].

In our system, graphs are built with arcs and two kinds of nodes, named s-entities and s-links. A s-entity models a spatial object of the choremes while a s-link models a spatial relation. The s-entities and the s-links can be qualified by different attributes. The arcs connect s-entities to s-links. They are labelled with the description of the roles of the spatial objects in the relations. For example, Figure 2 shows a part of the graph associated to the choreme of Figure 1: $\texttt{wood}_1$, $\texttt{meadow}_1$, $\texttt{cereal}_1$ are s-entities, $\texttt{isolate\&border}_1$ is a s-link, `isolating` and `isolated` denote the roles of the s-entities into the s-link $\texttt{isolate\&border}_1$. The s-link $\texttt{isolate\&border}_1$ means that $\texttt{meadow}_1$ is between but borders $\texttt{wood}_1$ and $\texttt{cereal}_1$, and that $\texttt{meadow}_1$ has a 'buffer' function: the cereal field is isolated from the wood thanks to the meadow.

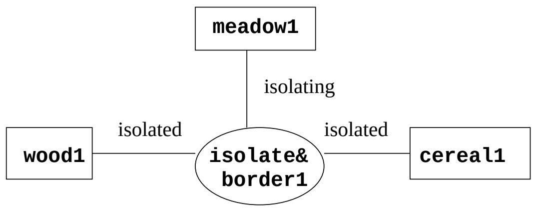

Figure 2. Part of the graph associated to the choreme of figure 1 (graph A).

2 Reasoning principles

The purpose of our system is to compare the spatial organization and functioning of farms. The farms previously studied are stored in a case base. The analysis of a new farm is based on the case-based reasoning (CBR) paradigm, that is, it uses past experiences (called cases) to solve new problems [9]. CBR is based on three main operations: retrieval, adaptation and storage. The goal of retrieval is to find a case `srce` in the case base similar to the target problem `tgt`. Adaptation uses the retrieved case `srce` in order to build a solution `Sol(tgt)` to `tgt`. If the new case `(tgt, Sol(tgt))` is of interest, it is stored in the case base. Similarity between cases is defined according to domain knowledge and according to the characteristics of the current problem.

At the moment, we have worked on the retrieval operation and on case representation. In our system, a case is a graph representing a spatial organization associated with a

[1] Actually choremes are basic graphical units used to symbolize geographical patterns. The maps are based on choremes, but the agronomists call them *choremes*.

functional explanation. For instance the graph A of Figure 2 and the following explanation *"the meadow isolates crop fields from the wood and thus protects crops from dampness and shade due to the wood"* constitute a case.

We propose to define the similarity between two cases using graph properties and a hierarchical representation of domain knowledge. In our system, domain knowledge includes agronomic concepts and qualitative spatial relations and their inter-relations. A part of the hierarchy used is described in Figure 3. The spatial relations are those used by the agronomists to describe spatial organization (e.g. separate, isolate, near, far, cross over, etc.), i.e. they include spatial and functional aspects. They are organized according to the usual spatial relations proposed in the QSR domain [11, 5, 12]. For instance, SEPARATE and ISOLATE are considered to be sort of IS-BETWEEN relation with different functional meanings. The relation ISOLATE&BORDER, as said before, is a combination of the relations BORDER and ISOLATE. The relation BORDER is itself a kind of the topological relation EC, where the regions share a line rather than a single point.

The cases and the concept hierarchy are implemented with the description logic system RACER [4] which is very expressive and capable of dealing with concepts as well as with individuals. In particular, the classification mechanism of RACER is used for graph comparison as shown below.

3 Similarity between graphs

The similarity between two graphs is computed with the help of an *edition distance* that relies on three operations, deletion, addition or substitution of nodes, as proposed in [1]. The substitution of nodes relies on a classification mechanism acting upon the concept hierarchies. The deletion or addition of nodes rely on specific rules. The distance between two graphs is evaluated on the basis of the type and number of operations needed to transform a graph into another.

Let us take an example. Suppose we want to compare the graph A of Figure 2 with the graph B described in Figure 4, that represents a field of cereal bordered by a wood.

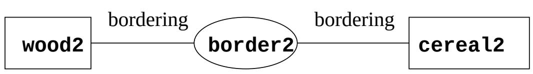

Figure 4. Another graph (B).

First the nodes of the two graphs are classified according to the hierarchy of Figure 3. The node $\texttt{meadow}_1$ of graph A is classified into the concept MEADOW, while the nodes $\texttt{cereal}_1$ and $\texttt{cereal}_2$ of graph B are classified into the concept CEREAL. The nodes $\texttt{wood}_1$ and $\texttt{wood}_2$ are classified into the concept WOOD. The node $\texttt{border}_2$ is classified into the concept BORDER and the node $\texttt{isolate\&border}_1$ is classified into the concept ISOLATE&BORDER that is a sub-concept of IS-BETWEEN.

The next step is to match the nodes of the two graphs: $\texttt{cereal}_1$ and $\texttt{cereal}_2$ can be matched as they belong to the same concept; $\texttt{wood}_1$ and $\texttt{wood}_2$ can be matched for the same reason; $\texttt{meadow}_1$ could be matched with $\texttt{cereal}_2$ as they belong to the same super-concept AGRICULTURAL-LAND, but the matching of $\texttt{cereal}_1$ and $\texttt{cereal}_2$ is better; on the contrary, $\texttt{wood}_1$ and $\texttt{cereal}_2$ cannot be matched as woods and cereals have completely different function in the farms considered. The matchings are evaluated according to the distance between concepts w.r.t the hierarchy.

Finally, the problem is to match the s-link $\texttt{border}_2$ that represents a binay relation with the s-link $\texttt{isolate\&border}_1$ that represents a ternary relation. This is done as follows. First the nodes $\texttt{meadow}_1$ and $\texttt{cereal}_1$ can be classified into the same super-concept AGRICULTURAL-LAND. Furthermore, as the relation ISOLATE&BORDER is a kind of IS-BETWEEN relation, the following rule can be applied: *if a region b is between and borders two regions a and c, and b can be assimilated to a or c, then the two regions a and c border each other*. Thus, the two nodes $\texttt{meadow}_1$ and $\texttt{cereal}_1$ are substituted by only one node $\texttt{agri-land}_1$, the node $\texttt{isolate\&border}_1$ is deleted and a new node $\texttt{border}_1$ is created that connects $\texttt{agri-land}_1$ and $\texttt{wood}_1$. Finally the graph A is transformed into the new graph A' (Figure 5).

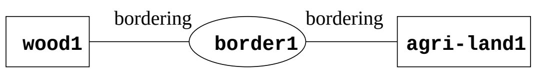

Figure 5. The new graph A' transformed from A.

In graph B, the node $\texttt{cereal}_2$ can be classified into the super-concept AGRICULTURAL-LAND and thus can be substituted with a node $\texttt{agri-land}_2$. The resulting graph B' is identical to graph A'. The distance between A and B can be evaluated according to the operations used to transform A into A' and B into B', that is a substitution (of two nodes by one node), a deletion and a creation on the one hand, and a substitution on the other hand. The sequence of the operations used to transform a graph into another graph defines a *similarity path*, that can be used in CBR for the adaptation operation [10]. In our example, the explanation associated to graph A should be adapted into a new explanation associated to graph B, e.g., *"crop fields are not isolated from the wood and thus crops are not protected from dampness and shade due to the wood"*.

It should be noticed that a graph C representing a wood bordering a meadow can be compared with B with fewer transformations and can lead to another explanation. A major problem is thus to choose the good matching between graphs or to combine the results of several matchings.

Summary

This paper presents a preliminary work done on the design of a case-base explanation system for helping agronomists analyze the relationship between farm spatial organization and functioning. In our system, a case is a graph representing (part of) a farm spatial organization associated with a functional explanation. The reasoning principle is to use the past cases to help the analysis of a new case. The retrieval operation is based on graph similarity that is computed with the help of an edition distance.

Graphs are described with agronomic spatial objects and spatial and functional relations. The matching between two graphs relies on domain knowledge and on spatial inferences. This paper gives an example of the knowledge and inferences used, but there is still a lot of work to be done, especially to define general rules for matching spatial structures and then for adapting functional explanations.

REFERENCES

[1] H. Bunke and B.T. Messmer, 'Similarity Measures for Structured Representations', in *Proceedings of the First European Workshop on Case-Based Reasoning (EWCBR'93), Kaiserslautern*, eds., M.M. Richter, S. Wess, K.-D. Althoff, and F. Maurer, pp. 26–31, (1993).

[2] M. Capitaine, 'Graphical representation of farm spatial organisation in Lorraine'. INRA Internal Document, 2001.

[3] M. Chein and M.-L. Mugnier, 'Conceptual graphs: Fundamental notions', *Revue d'Intelligence Artificielle*, **6**(4), 365–406, (1992).

[4] V. Haarslev, R. Möller, and A.-Y. Turhan. Racer user's guide and reference manual. Hamburg Universität, 2001.

[5] D. Hernández, *Qualitative Representation of Spatial Knowledge*, LNAI 804, Springer-Verlag, 1994.

[6] S. Lardon, C. Capitaine, and M. Benoît, 'Les modèles graphiques pour représenter l'organisation spatiale des activités agricoles.', in *Représentations graphiques dans les systèmes complexes naturels et artificiels. Journées de Rochebrune*, pp. 127–150. ENST, (2000).

[7] F. Le Ber, L. Mangelinck, and A. Napoli. Design and comparison of lattices of topological relations for spatial representation and reasoning. INRIA Research Report RR-4321, 2001. 57 pages.

[8] F. Le Ber and A. Napoli, 'Representation of spatial relations and structures in object-based knowledge representation systems', in *ECAI Workshop on Current Issues in Spatio-Temporal Reasoning*, ed., R. V. Rodríguez, pp. 81–86, Berlin, (2000).

[9] *Case-Based Reasoning. Experiences, Lessons, & Future Directions*, ed., D. B. Leake, AAAI Press / The MIT Press, 1996.

[10] J. Lieber and A. Napoli, 'Correct and Complete Retrieval for Case-Based Problem-Solving', in *Proceedings of the 13th European Conference on Artificial Intelligence (ECAI'98), Brighton, UK*, ed., H. Prade, pp. 68–72. John Wiley & Sons Ltd., (1998).

[11] D. A. Randell, Z. Cui, and A. G. Cohn, 'A Spatial Logic based on Regions and Connection', in *3rd International Conference on Principles of Knowledge Representation and Reasoning*, pp. 165–176. Morgan Kaufmann Publishers, (1992).

[12] L. Vieu, 'Spatial Representation and Reasoning in Artificial Intelligence', in *Spatial and Temporal Reasoning*, ed., O. Stock, 5–42, Kluwer Academic Publishers, (1997).

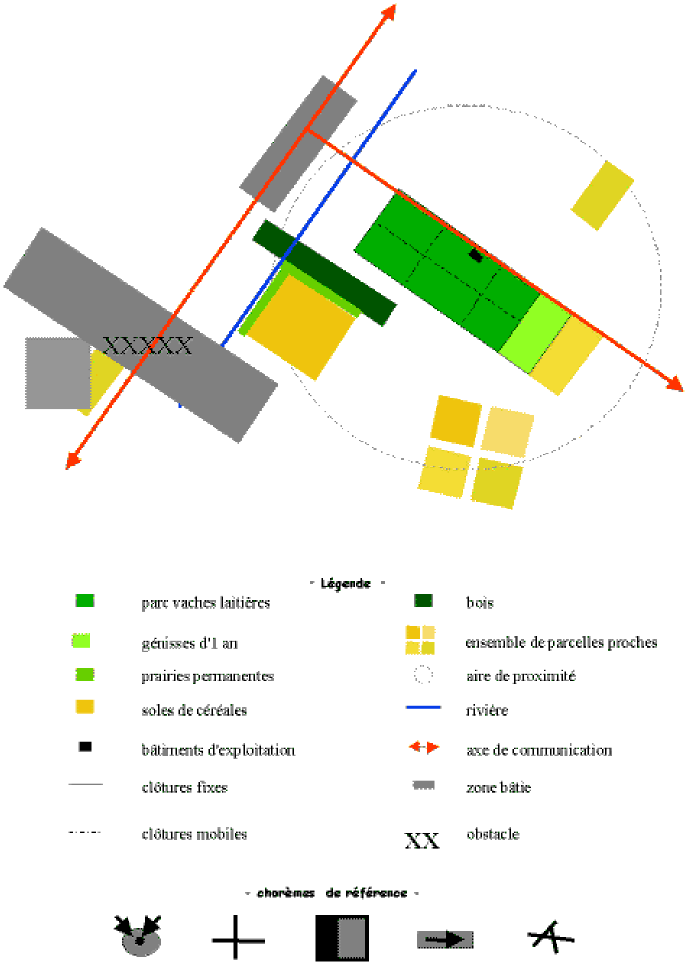

Figure 1. A choreme representing the spatial organization and functioning of a farm in Lorraine (east of France) [2].

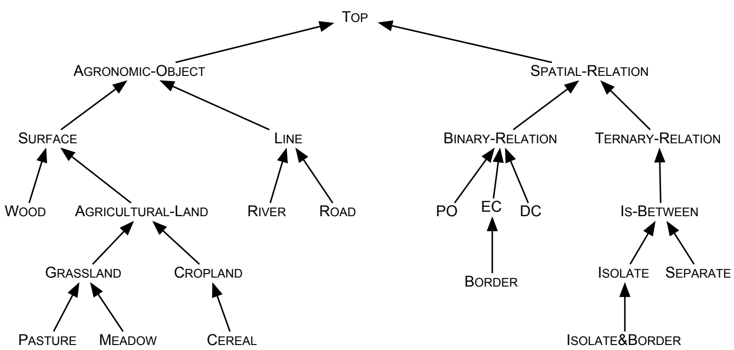

Figure 3. Part of the concept hierarchy designed for the representation of agronomic knowledge and spatial relations.